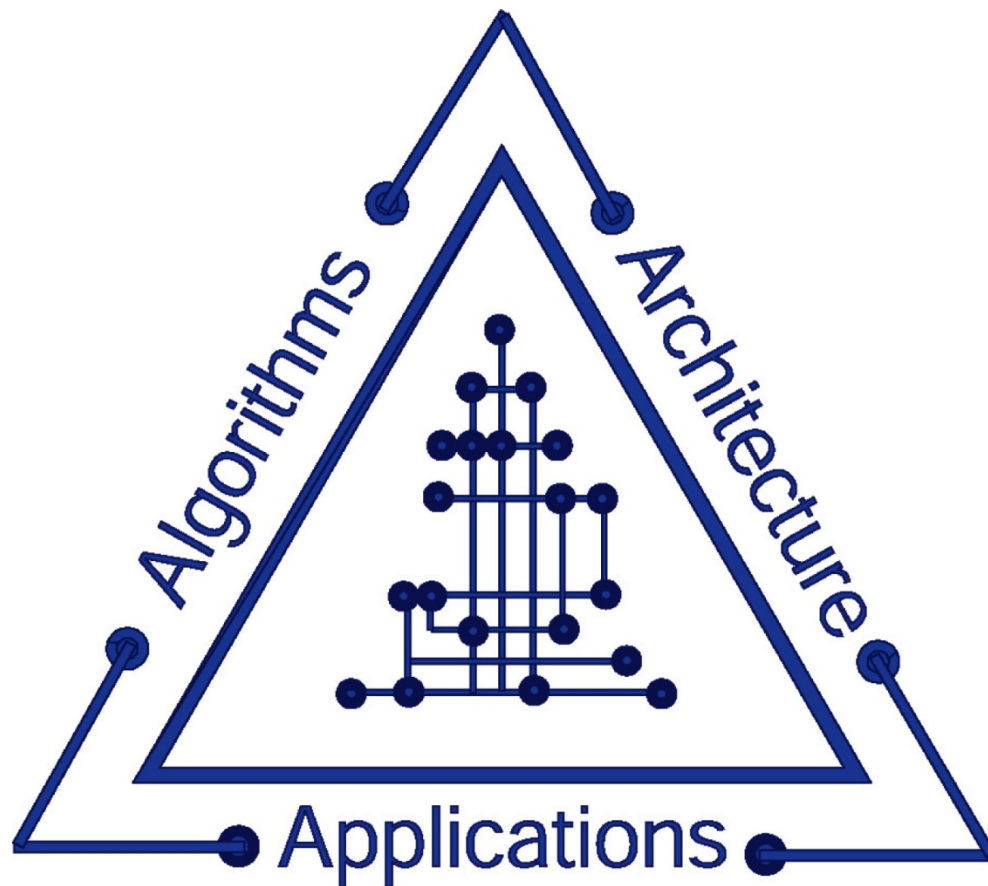


The Salishan Conference on HIGH-SPEED COMPUTING

LANL / LLNL / SNL



April 27 - 30, 2015

*Salishan Lodge
Gleneden Beach, Oregon*

Welcome

Welcome to the Salishan Conference on High-Speed Computing. This conference was founded in 1981 as a means of gathering experts in computer architecture, languages, and algorithms together to improve communication, develop collaborations, solve problems of mutual interest, and provide effective leadership in the field of high-speed computing. Attendance at the conference is by invitation; we limit attendance to about 150 of the world's brightest people. Attendees are from national laboratories, academia, government, and private industry. We keep the conference small to preserve the level of interaction and discussion among the attendees.

The conference agenda and selection of participants has been designed to focus discussion on technical issues of relevance to our conference theme: Resilience at Scale. The talks have been selected to give attendees information about the latest technologies and issues facing high-speed computing. The evening sessions are structured to encourage informal discussions and networking among all of the participants.

If you have any comments or suggestions for future topics and/or speakers, we encourage you to speak to any of the conference committee members.

We hope you find this conference stimulating, challenging, and also relaxing—enjoy!

Conference Committee

Brian Carnes & Bert Still, *LLNL*
Manuel Vigil & Allen McPherson, *LANL*
Jim Ang & Ron Brightwell, *SNL*

Logistics

Conference sessions and the Random Access session will be held in the Long House. Lunches and the working dinner will be held in the Council House.

For administrative support, please speak to Jan Susco, Dee Cadena or Suz Espinoza, located in the registration area (Salal Room). If you have specific questions regarding audiovisual equipment or network connectivity, please seek out administrative support.

Next Conference Dates: April 25-28, 2016
 April 24-27, 2017

MAIN LODGE MAP

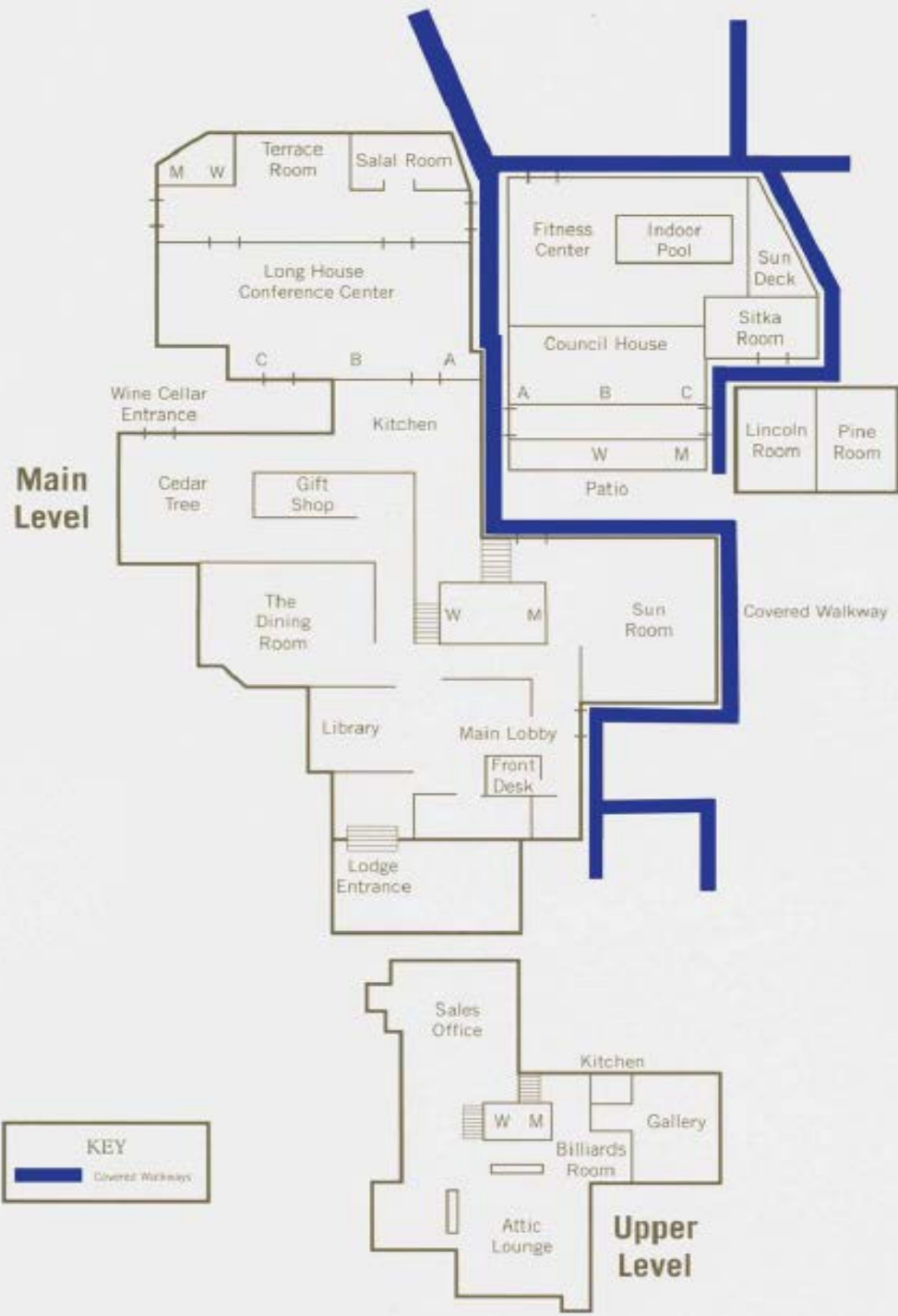


Table of Contents

The Conference on High-Speed Computing

Welcome and Logistics	2
Lodge Map	3
Sponsorship	6
Conference Theme	8
Conference Program		
Monday:	Keynote.....	10
Tuesday:	Session 1: The Fault Environment.....	11
	Session 2: Resilient Numerical Methods.....	12
	Working Dinner.....	13
Wednesday:	Session 3: System Software and APIs.....	14
	Random Access.....	15
	Student Poster Session.....	15
Thursday:	Session 4: Data Analysis on Uncertain Data.....	16
	Session 5: Future Application Development Environment...17	
Abstracts	18
Attendees	33
Conference Notes	41

THIS PAGE LEFT BLANK
INTENTIONALLY

Sponsorship

The Salishan Conference on High-Speed Computing is organized and hosted by Lawrence Livermore, Los Alamos and Sandia National Laboratories. Additional sponsorship for the evening portions of our program is provided by the corporations listed here. The Association for High-Speed Computing (AHSC) is the event sponsor and is delighted to co-organize with LANL, LLNL and SNL the 2015 Salishan Conference on High-Speed Computing.

One of the highlights of the conference is the informal discussions held each evening. These sessions help us to go beyond the formal presentations to exchange ideas, solve problems and develop friendships.

This year the following companies are helping to sponsor the evening sessions:

Advanced Micro Devices, Inc.

ARM

Cray, Inc.

DataDirect Networks, Inc.

EMC

Hewlett-Packard Company

IBM Corporation

Intel Corporation

Micron Technology, Inc.

NVIDIA Corporation

PGI Compilers & Tools

Rogue Wave Software, Inc.

Silicon Graphics International Corp.

We would like to express our thanks to these companies for their generous support.

THIS PAGE LEFT BLANK
INTENTIONALLY

Conference Theme

Resilience at Scale

Over the past 5-6 years, there has been significant research, studies, and several workshops related to resilience at scale and significant progress has been made in many related domains. Despite this progress, the exascale resilience problem is not fully understood, and the community is still facing the difficult challenge of ensuring that exascale applications complete and generate correct results while running on unreliable systems. System developers and researchers are working to characterize and quantify the types and frequencies of known faults and methods to mitigate them. However, it is not clear that this knowledge has been effectively passed to applications developers or that they fully understand the broad range of issues that a fault-prone environment will bring to the development and usage of their codes.

Developers of HPC codes face two key questions: What are the reliability requirements for their particular code and how do they construct it and execute it to meet these requirements? Answers to the first question can largely be determined by analysis of use cases based on expected workload and workflow. Finding answers to the second question is much more difficult and requires deep and broad knowledge of the expected fault environment as well as the landscape of technology and methods available to respond to and manage these faults. Novel numerical methods or more stochastic approaches may be required to meet accuracy requirements in the face of undetectable soft errors. Developers might use standard fault notification APIs and adaptive system runtimes to respond to node failures. Burst buffers can provide a solution for in-memory data replication and automated recovery from crashes. In any case, application developers must be aware of these potential solutions and have the ability to feed their requirement to the developers of these products.

The goal of this year's Salishan conference is to expose the audience to the broad range of resilience issues that will affect the HPC community. Invited talks will focus on recent research in areas that are particularly important to the development of resilient applications on exascale systems. Application developers will learn about the diverse fault environments they can expect to face in the near future. Hardware providers, algorithm developers, and system software and library developers will have ample opportunity to hear directly from application developers on what they expect and can tolerate regarding resiliency at scale. The main conference goal, as always, is to provide ample forums for discussion among participants to provide feedback, discuss issues, explain concerns, develop collaborations and recommend solutions. This year the conference organizing committee will also seek to foster increased participation with younger staff, women and minorities.

Session 1: The Fault Environment

This session is designed to expose the audience to the faults we expect to face as systems move to exascale. What components of the system are expected to fail and at what rate? What is the nature of those failures? What failures can we detect and what is the cost of detection to applications? What are the potential effects of failures that we can't detect? What does it really mean for a large-scale system to be "reliable"? Is this fault environment expected to improve or worsen in the future and to what extent? What are the expected trade-offs between cost and system reliability? What can we learn from the private sector and their approaches to reliability?

Session 2: Resilient Numerical Methods

This session will deal with potential approaches that deal with soft errors, both detectable and undetectable, in machine logic. How can error and uncertainty be quantified? Is redundant computation a reasonable approach? How do stochastic methods compare with deterministic methods in the presence of undetected soft errors? Can compilers help with soft error mitigation? To what extent will we sacrifice flexibility with known-performance standards? What are applications currently doing to mitigate faults and what are the costs?

Session 3: System Software and APIs

This session will expose the application community to current research in system software that is designed to support resilient applications on large-scale systems. How will system software enable resilience? How transparent will it be to applications? What is the expected cost to applications (data redundancy, etc.)? Has progress been made toward development of a standardized fault-handling model? How will system software serve as the interface between applications and the underlying hardware fault environment?

Session 4: Data Analysis on Uncertain Data

This session examines methods and techniques to deal with potentially corrupt or unreliable data (produced by unreliable machines). Can data analysis tools provide a means to detect unreliable data? To quantify the uncertainty? What additional information can the system supply (detected failures, average rate of undetected failures, etc.) to help analysis software predict the reliability of data?

Session 5: Future Application Development Environment

This session deals with how the Application Development Environment might change in the face of a constant-fault environment. Will new languages emerge with features that enable response to faults? Will large-scale databases re-emerge as data caches for continuous restart? Will we move to a more dynamic computing environment with codes running concurrently and elastically with other codes? At what level do application developers want to deal with failure (full system support, application notification and response, etc.)? How important is portability? What are the most important tools to help application developers deal with large scale system failures?

Conference Program

Resilience at Scale

Monday, April 27, 2015

4:30–7:00 pm

**Registration
(Salal Room)**

6:00 pm

Welcome/Keynote Address

**Title: Failure, Resilience, Opportunity
 and Innovation**

Speaker: John Daly, *Department of Defense*

8:00 pm

**Reception and Informal Discussions
(Immediately following the Keynote in the
Council House)**

Tuesday, April 28, 2015

8:00 am **Registration Opens (Salal Room)**

Breakfast available (Terrace)

8:30 am **Session 1: The Fault Environment**

Title: **Building Reliable Chips in Future
Technologies: Fact, Fiction or An Oxymoron?**

Speaker: Vikas Chandra, *ARM*

Title: **The Fault Environment Unveiled**

Speaker: Sudhanva Gurumurthi, *AMD/University of Virginia*

Title: **Resiliency for Reliability – Myths and Truths**

Speaker: Shekhar Borkar, *Intel Corporation*

10:00 am **Break**

Refreshments available (Terrace)

10:30 am **Title:** **New Resilience Capabilities with
Micron's HMC**

Speaker: David Resnick, *Sandia National Laboratories*

11:00 am **Panel Discussion**

Tuesday, April 28, 2015

12:00 pm	Lunch (Council House)
1:30 pm	Session 2: Resilient Numerical Methods
	Title: Bend but Don't Break: Prospects for Resilience without Recovery in Algorithms for Hyperbolic Systems
	Speaker: Jeffrey Hittinger, <i>Lawrence Livermore National Laboratory</i>
	Title: Fault Tolerance in Numerical Library Routines
	Speaker: Jack Dongarra, <i>University of Tennessee</i>
	Title: On Numerical Resiliency in Numerical Linear Algebra Solvers
	Speaker: Luc Giraud, <i>Inria</i>
3:00 pm	Break
	Refreshments available (Terrace)
3:30 pm	Title: Application Structure Aware Resiliency and Cost Model for Differentiated Recovery
	Speaker: Anshu Dubey, <i>Lawrence Berkeley National Laboratory</i>
4:00 pm	Panel Discussion

Tuesday, April 28, 2015

6:00 pm

Working Dinner/Speaker (Council House)

Title: Why HPC Matters

Speaker: Eng Lim Goh, *Silicon Graphics International Corp.*

8:00 pm

**Reception and Informal Discussions
(Immediately following the Working Dinner in the
Cedar Tree Room)**

Wednesday, April 29, 2015

8:00 am Introduction to Sessions

Breakfast available (Terrace)

8:30 am Session 3: System Software and APIs

Title: Data-Driven Decision Making in Resilience

Speaker: Nathan DeBardeleben, *Los Alamos National Laboratory*

Title: Revisiting Checkpointing for Exascale-Class Systems

Speaker: Kurt Ferreira, *Sandia National Laboratories*

Title: Scalable Program Analyses to Improve Software Reliability

Speaker: Cindy Rubio-Gonzalez, *University of California at Davis*

10:00 am Break

Refreshments available (Terrace)

10:30 am Title: Fault Tolerant Programming Abstractions and Failure Recovery Models for MPI Applications

Speaker: Ignacio Laguna Peralta, *Lawrence Livermore National Laboratory*

11:00 am Panel Discussion

Wednesday, April 29, 2015

- 12:00 pm** **Lunch on your own**
- 1:30 pm** **No Scheduled Session**
- 5:00 pm** **Random Access (Long House)**

The Random Access session consists of timely communications from participants on areas of interest to the Conference. Presentations are strictly limited to 10 minutes. A sign-up board is provided in the registration area.

- 8:00 pm** **Reception and Informal Discussions (Immediately following Random Access in the Council House)**

Student Poster Session (Immediately following Random Access in the Council House)

This conference hosts and selects students from each of the sponsoring laboratories, inviting them to present posters and discuss their research with our Salishan community. All conference attendees are encouraged to visit with this year's students:

Brian Atkinson, *Clemson University*

Jon Calhoun, *University of Illinois*

Ken Czuprynski, *University of Iowa*

Ana Gainaru, *University of Illinois at Urbana-Champaign*

Nalini Kumar, *University of Florida*

Aaditya Landge, *University of Utah*

Thursday, April 30, 2015

8:00 am **Introduction to Sessions**

Breakfast available (Terrace)

8:30 am **Session 4: Data Analysis on Uncertain Data**

Title: **Relaxing Resilience Data Quality Requirements Due to Visualization and Analysis Needs**

Speaker: James Ahrens, *Los Alamos National Laboratory*

Title: **Living With "Dirty" Data While Avoiding Exascale "Garbage In, Garbage Out"**

Speaker: Michael McKerns, *California Institute of Technology*

Title: **Approximate Computing for Approximate Data**

Speaker: Martin Rinard, *Massachusetts Institute of Technology*

10:00 am **Break**

Refreshments available (Terrace)

10:30 am Title: Towards Interactive Analysis and Exploration of the HPC Performance Landscape

Speaker: Yarden Livnat, *University of Utah*

11:00 am **Panel Discussion**

Thursday, April 30, 2015

12:00 pm	Lunch (Council House)
1:30 pm	Session 5: Future Application Development Environment
	Title: Three Crazy Ways to Cope with Failure That Will Change Your Apps Forever
	Speaker: Sung-Eun Choi, <i>Cray, Inc.</i>
	Title: Quantitatively Modeling Application Resilience with the Data Vulnerability Factor
	Speaker: Jeffrey Vetter, <i>Georgia Tech</i>
	Title: Global View Resilience: Flexible, Portable, Scalable Application Recovery for Fail-Stop and “Silent” Errors
	Speaker: Andrew Chien, <i>University of Chicago</i>
3:00 pm	Break
	Refreshments available (Terrace)
3:30 pm	Title: Exploiting the User’s Knowledge of Resilience
	Speaker: Robert Lucas, <i>Information Sciences Institute</i>
4:00 pm	Panel Discussion
6:00 pm	Reception and Informal Discussions (Council House)

Abstracts

Keynote Address

Failure, Resilience, Opportunity and Innovation

John Daly, *Department of Defense*

Thomas Edison once claimed, “I failed my way to success.” This may very well be true. His carbon filament incandescent light bulb emerged after thousands of unsuccessful attempts. What about high performance computing (HPC) though? Can we fail our way to success? HPC recognizes that traditional notions of supercomputer reliability cannot be expected to scale. In CMOS technology, process sizes are shrinking. In system architecture, component counts are increasing. In software, codes are becoming more complex. In an age of shrinking budgets, “more money” is not the solution to every problem. How will HPC continue to provide insight into the nation’s most important and challenging problems using computers that fail regularly and even give wrong answers? Resilience is not about making all of the errors go away. On the contrary, systems intended to run without errors often fail in the most catastrophic ways. Resilience is about understanding how systems fail and creating applications that can fail their way to success. Over the past decade the HPC research community has learned much, but there is still more to be done. How much and what kinds of “correctness” are required to perform the predictive science or analytic discovery mission? Resilience is an opportunity to think carefully and creatively about these issues. Resilience is a call to innovation in HPC.

Session 1: The Fault Environment

Building Reliable Chips in Future Technologies: Fact, Fiction or an Oxymoron?

Vikas Chandra, *ARM*

With nearly three decades of continued CMOS scaling, the devices have now been pushed to their physical and reliability limits. At these dimensions, the devices become quite fluid and fragile thus resulting in various kinds of unreliable behavior during their lifetime. Rapidly shrinking technology nodes and aggressive voltage scaling have increased the probability of single event upsets (SEU). With technology scaling, SEU is becoming more challenging due to decreasing device dimensions, node capacitance and aggressive supply voltage scaling. Additionally, the flip-flop SEU vulnerability is now getting close to that of SRAMs. Also, scaling to sub-20nm technology nodes changes the nature of reliability effects from abrupt functional problems to progressive degradation of the performance characteristics of devices and system components. The impact of unreliability results in time-dependent variability, directly translating into design uncertainty in manufactured chips. We have analyzed aging effects on various design hierarchies of Cortex R4 in 28nm running real-world applications. We have also quantified the dependencies of aging effects on switching-activity and power-state of workloads. Implementation results show that processor timing degradation can vary from 2% to 11%, depending on the workload.

The Fault Environment Unveiled

Sudhanva Gurumurthi, *AMD/University of Virginia*

The use of highly scaled technologies and large component counts pose significant reliability challenges for large-scale systems. Knowledge of failures that occur in such systems is valuable for driving design decisions for component and system vendors, as well as for the operators of those systems in the field. Field studies play a key role in gaining such knowledge. This talk will present data and insights gained through failure data analysis from supercomputers and cloud data centers to highlight the value of such studies, discuss implications for future exascale systems, and identify challenges that need to be addressed.

Session 1: The Fault Environment

Resiliency for Reliability – Myths and Truths

Shekhar Borkar, Intel Corporation

High performance computing performance increases by three orders of magnitude every decade, made possible by continued advances in technology, system architecture, and the software stack. The trend will continue, albeit with several daunting challenges, and system reliability being one of the foremost. Traditional approaches to provide reliability, such as tri-modular redundancy, will be prohibitive due to energy and dollar cost. Resiliency can provide the same reliability with only a fraction of the energy and dollar cost; however, it needs discipline in system design with close attention to HW/SW co-design. The community has taken resiliency to heart, but at times with arbitrary and haphazard approaches.

This talk will start with defining resiliency and its goals, discuss types of faults and the errors they cause, their behavior and probability, impact to the system, and build a generalized resiliency framework. Then we will discuss how to implement such a framework with HW/SW co-design, identifying HW and SW components ensuring that it does not impact system performance and cost. Finally, we will discuss potential recovery (checkpointing) schemes to gracefully recover from the faults. Such a holistic approach will provide the necessary system reliability, covering all types of known faults, at a fraction of the cost.

New Resilience Capabilities with Micron's HMC

David Resnick, *Sandia National Laboratories*

New memory technology from Micron enables features to be added to the system memory such that most all memory failures can be fully recovered, including errors of the interconnection paths, within memory modules that use those components. In addition, the mechanisms used to enable such resilience can be used such that sets of the memory modules can see whole modules fail and system operation continue.

And, as expect that the memory components are unfamiliar to most of the audience, the presentation includes information about the new memory components and the benefits that they enable, particularly for large systems, that have not been possible previously, and that can allow systems to be built that meet the goals of exascale HPC.

Session 2: Resilient Numerical Methods

Bend but Don't Break: Prospects for Resilience without Recovery in the Algorithms for Hyperbolic Systems

Jeffrey Hittinger, *Lawrence Berkeley National Laboratory*

Hard and soft errors are expected to occur more frequently on exascale-class supercomputers, and the common treatment of using global checkpoint-restart will be infeasible at these scales. It is likely that detection and recovery mechanisms will be built into many levels of future software stacks, but like many general solutions, these methods cannot take advantage of efficiencies that derive from special properties of the problem being solved. Particularly for silent data corruptions, the algorithms in the application may possess properties that mitigate the need for lower-level intervention. Such is the case with certain iterative algorithms that still converge (with one or more additional iterations) despite the occurrence of silent errors.

However, for the class of problems described by systems of hyperbolic partial differential equations (PDEs), which includes shock hydrodynamics, most algorithms are direct, not iterative. Furthermore, the fundamental nature of the solutions to hyperbolic problems is to preserve and propagate signals, not to damp them. On the other hand, the very nature of PDE-based simulation is approximation with controlled error, and the fact that one can locally trade accuracy for robustness provides a direction for the development of efficient fault-tolerant direct methods. The shock-capturing algorithms in use today contain several features that may be repurposed to make the algorithm itself robust to silent data corruption. In this talk, I will decompose the standard flux divergence update algorithm into components for which suitable fault-tolerant mechanisms can be identified and provide preliminary results that show the promise of this approach. Based on these results, I will discuss likely directions for fruitful future research in this area.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. LLNL-ABS-668001.

Session 2: Resilient Numerical Methods

Fault Tolerance in Numerical Library Routines

Jack Dongarra, *University of Tennessee*

On the way to Exascale a generic, low-overhead, resilient extension becomes a desired feature of any programming paradigm. We explore a number of approaches to a dynamic task-based runtime using Algorithm Based Fault Tolerance for our numerical linear algebra libraries to build a generic framework providing soft error resilience to task-based programming paradigms. The first recovers the application by re-executing the minimum required sub-DAG, the second takes critical checkpoints of the data flowing between tasks to minimize the necessary re-execution, while the last one takes advantage of algorithmic properties to recover the data without re-execution. These mechanisms have been implemented in the PaRSEC task-based runtime framework. Experimental results will be shown for our approach and quantify the overhead introduced by such mechanisms.

On Numerical Resiliency in Numerical Linear Algebra Solvers

Luc Giraud, *Inria*

In this talk we will discuss possible numerical remedies to survive data loss in some numerical linear algebra solvers namely Krylov subspace linear solvers and some widely used eigensolvers. We will present a new class of numerical fault tolerance algorithms at application level that does not require extra resources, i.e., computational unit or computing time, when no fault occurs. Assuming that a separate mechanism ensures fault detection, we propose numerical algorithms to extract relevant information from available data after a fault. After data extraction, well-chosen part of missing data is regenerated through interpolation strategies to constitute meaningful inputs to numerically restart the algorithm. We have designed these methods called interpolation-restart techniques for the solution of linear systems and eigensolvers. We will also present some preliminary investigations to address soft error detection again at the application level in the conjugate gradient framework.

Finally, we will expose the numerous open questions that we are facing that hopefully will lead to fruitful discussions.

This is a joint research effort with E. Agullo (Inria), E.F. Yetkin (Inria), P. Salas (Sherbrooke Univ.) and M. Zounon (Inria).

Session 2: Resilient Numerical Methods

Application Structure Aware Resiliency and Cost Model for Differentiated Recovery

Anshu Dubey, *Lawrence Berkeley National Laboratory*

In order to achieve resiliency by saving state, applications necessarily incur overhead. The cost of recovery is a combination of overheads, reconfiguration of state from the saved one, and recomputing to get back to the intermediate state where the loss occurred. Therefore, a cost model that can evaluate the trade-offs becomes an extremely important part of the resiliency strategy for any application, especially multi-scale and multi-component applications that can exploit differentiated state saving for different scales and components. We examine such a differentiated recovery approach and its corresponding cost model in the context of structure adaptive mesh refinement (SAMR). We expect that the methodology behind our approach would be applicable to many applications.

Dinner Speaker

Why HPC Matters

Eng Lim Goh, *Silicon Graphics International Corp.*

HPC may be daunting to the uninitiated, but it's touching ordinary folks in ways they may not even realize. HPC is playing a role in helping humanity solve its hardest problems. This presentation will shed light on some of the most profound HPC use cases of our time, spanning nearly every aspect of society and ranging from advanced manufacturing to disaster warning systems to improving care for cancer patients. Whether it's meeting basic needs, reducing hardships, promoting industry or answering the profound questions of the universe, HPC is there. At the same time as HPC grows into new markets and segments, the traditional application areas are as relevant and vital as ever. This presentation will cover the range of application areas and discuss how HPC contributes to them.

Session 3: System Software and APIs

Data-Driven Decision Making in Resilience

Nathan Debardeleben, *Los Alamos National Laboratory*

In this talk, Dr. Nathan DeBardeleben will discuss collaborative applied research conducted with vendors and other laboratories that looks at current behaviors of DOE supercomputers. Nathan will show how beneficial this data is and how it can lead to some surprising conclusions as well as give confidence in the systems we are running. He will use the data to look at projections for next generation systems and show how there are several paths forward that can lead to varying levels of concerns. Nathan will show how real data can lead to decision making that can inform procurement as well as production operation of systems of today. Finally, he will discuss how today's field data can help application scientists prepare for the challenge of next generation systems.

Revisiting Checkpointing for Exascale-Class Systems

Kurt Ferreira, *Sandia National Laboratories*

Checkpoint/restart has remained a popular method for addressing fault-tolerance on large-scale production HPC systems for the past decade. Its success can largely be attributed to the following properties: 1) It is application and platform independent; 2) Overheads remain low on current systems as the time to write a checkpoint and the time to restart is significantly smaller than the mean time between failure (MTBF); and, 3) Saved checkpoints contain no errors because silent data corruption (SDC) is an exceedingly rare event. There is concern, however, that these properties will no longer hold on future extreme-scale systems. In particular, shrinking MTBFs reducing efficiencies, shrinking feature sizes leading to frequent SDCs, and power capacity limits threaten the future of checkpointing. In this talk, we revisit the future of checkpointing at extreme-scale. Using failure data from current leadership-class systems, along with some healthy speculation on the makeup on future systems, we will attempt to address the question of whether a path exists to keep checkpoint/restart viable on future exascale-class systems.

Session 3: System Software and APIs

Scalable Program Analyses to Improve Software Reliability

Cindy Rubio-Gonzalez, *University of California at Davis*

In this talk, I will present my work on developing and applying static program analyses to automatically find error-propagation bugs in large software systems. Bugs found in error handlers are among the most pervasive, dangerous, and difficult to detect in software systems. Incorrect error handling is a longstanding problem in many application domains, including systems software and user applications that use the return-code idiom. First, I will give an overview of an interprocedural context- and flow-sensitive analysis that tracks the propagation of error codes. This analysis is formalized using weighted pushdown systems (WPDS). I will describe how the analysis is used to find a variety of error-propagation bugs, such as dropped errors, misused error-valued pointers, and error-code mismatches between source code and error-reporting program documentation. I will present results for numerous real-world Linux file systems such as ext3 and ReiserFS, and Linux device drivers where the analysis have found hundreds of confirmed error-propagation bugs. As for user applications, I will present results for the Mozilla Firefox web browser. Last, I will briefly talk about our most recent work on database-backed program analysis for scalable error propagation.

Fault Tolerant Programming Abstractions and Failure Recovery Models for MPI Applications

Ignacio Laguna Peralta, *Lawrence Livermore National Laboratory*

Resilience to faults remains a challenge as we push toward exascale computing. A large fraction of software and hardware faults become visible to users as process or node failures. These failures often show up at the message passing interface (MPI) layer, which is the dominant “glue” to build large-scale HPC applications. The MPI programming model, however, provides no resilience mechanisms to applications—if a process or node dies, applications can do little more than abort. In this talk, I will describe the problem of incorporating fault-tolerance abstractions into MPI applications and give an overview of possible solutions to the problem. I will also evaluate, via modeling and simulations, some of the current failure mitigation approaches for MPI, with an emphasis on their programmability and their performance in current (petascale) and future (exascale) systems. (LLNL-ABS-667810)

Session 4: Data Analysis on Uncertain Data

Relaxing Resilience Data Quality Requirements Due to Visualization and Analysis Needs

James Ahrens, *Los Alamos National Laboratory*

In this talk, I discuss the potential problem of system faults, and what affect they have on the overall quality of scientific results. Many simulation data analysis products are produced as part of parameterized ensembles and many of these data products are mathematically integrated simulation results over time and space. Exascale arrays of double precision floating point values are likely rendered to megapixel screens with eight to sixteen bit color resolution. Many traditional resilience approaches strive to reproduce the exact simulation results that would have occurred without the fault. In the context of how most simulation results will be used this is likely too strong of a requirement. We present examples of the resilience of visualization and analysis algorithms to approximate data errors. We believe that reproducibility requirements for resilience approaches can be relaxed, resulting in the improved application performance, while still providing quality scientific answers within a prescribed error bound.

Session 4: Data Analysis on Uncertain Data

Living With "Dirty" Data While Avoiding Exascale "Garbage In, Garbage Out"

Michael McKerns, *California Institute of Technology*

Large-scale calculations on exascale platforms are expected to have strongly probabilistic behavior. With computations of this size and complexity, the corruption and unreliability of data will not only occur, it will be commonplace.

Data uncertainty (in some form) may happen with enough frequency that probabilistic methods may even need to be adopted to verify the exascale system is operating as expected. While data resilience at scale may be a frontier in high-performance computing, fortunately the science of statistics on uncertain data is not new. We will look at statistical methods for analyzing data quality, including: outlier detection, sampling statistics, and significance testing, as well as building models of system behavior and validating data against the statistically estimated models.

The practical application of statistical methods to the exascale environment opens up many research questions. For example, how will these types of statistical tests and diagnostics impact system performance? Can we build algorithms that analyze distributed or non-homogenous data, and can we do this analysis without introducing new errors or further corrupting data? Can we implement data analysis and data validation algorithms that are efficient, accurate, and real-time? Alternatively, what algorithmic changes might we make to provide robustness against data uncertainty?

Session 4: Data Analysis on Uncertain Data

Approximate Computing for Approximate Data

Martin Rinard, *Massachusetts Institute of Technology*

Approximate data is a fact of life that the computing community has long dealt with. Approximate computing is a relatively new computing paradigm that gives up some accuracy in return for energy and/or performance savings. One may very reasonably ask why one should use an expensive exact computing framework to compute on approximate data. The answer is that such a combination is likely to be suboptimal and that considering both forms of uncertainty together is likely to generate a better end-to-end result. This talk will consider ways to consider the combined uncertainty associated with the approximate data and approximate computation together. It will also consider new, more flexible approximate computing models that work with larger classes of approximate hardware.

Towards Interactive Analysis and Exploration of the HPC Performance Landscape

Yarden Livnat, *University of Utah*

The evolution of high performance computing platforms towards the goal of exascale computing is delivering ever more powerful supercomputers. However, exploiting the full capabilities of these machines is becoming exponentially more difficult with each new generation of hardware because of its complexity and heterogeneity. In this talk I will present recent advances in analyzing and visualizing large scale performance information. In particular, I will focus on the need for interactive exploratory tools and the challenge that they pose given the size and diversity of the performance information. I will describe a few success stories and discuss current and upcoming challenges that emerge in dealing with incomplete information in the moving target new hardware components that can only be represented as a high dimensional space of performance scenarios.

Session 5: Future Application Development Environment

Three Crazy Ways to Cope with Failure That Will Change Your Apps Forever

Sung-Eun Choi, *Cray, Inc.*

In an ideal world application scientists would think deep math thoughts, supercomputers would blissfully compute and great science would be done. In the real world failures abound and computing is anything but blissful. With perfect reliability out of reach and getting more so, what can we do? In this talk, I will present some ways scientists and software designers can cooperate to achieve resilience in the face of failures.

Quantitatively Modeling Application Resilience with the Data Vulnerability Factor

Jeffrey Vetter, *Georgia Tech*

Recent strategies to improve the observable resilience of applications require the ability to classify vulnerabilities of individual components (e.g., data structures, instructions) of an application, and then, selectively apply protection mechanisms to its critical components. To facilitate this vulnerability classification, it is important to have accurate, quantitative techniques that can be applied uniformly and automatically across real-world applications. Traditional methods cannot effectively quantify vulnerability, because they lack a holistic view to examine system resilience, and come with prohibitive evaluation costs. To attack this problem, we introduce a data-driven, practical methodology to analyze these application vulnerabilities using a novel resilience metric: the data vulnerability factor (DVF). DVF integrates knowledge from both the application and target hardware into the calculation. To calculate DVF, we extend the Aspen performance modeling language to provide a structured, fast modeling solution. We evaluate our methodology on six representative computational kernels; we demonstrate the significance of DVF by quantifying the impact of algorithm optimization on vulnerability, and by quantifying the effectiveness of specific hardware protection mechanisms. This is joint work with Jeremy Meredith, Sparsh Mittal, Li Yu, and Dong Li.

Session 5: Future Application Development Environment

Global View Resilience: Flexible, Portable, Scalable Application Recovery for Fail-Stop and “Silent” Errors

Andrew Chien, *University of Chicago*

Global View Resilience (GVR) is a library for application-controlled error recovery. GVR provides portable, scalable resilience and has been demonstrated with real applications at 16,384 nodes on Cray XC30 and IBM BG/P systems. Beyond rollback-restart recovery, GVR enables several new dimensions of flexibility for applications to manage data redundancy and employ novel recovery methods that exploit deep science and code knowledge to achieve high efficiency and resilience.

Specifically, GVR begins with a subset of the Global Arrays interface, and adds new capabilities to create versions, name versions, and compute on version data. Applications can focus versioning where and when it is most productive, and customize for each application structure independently. This control is portable, and its embedding in application source makes it natural to express and easy to maintain. The ability to name multiple versions and “partially materialize” them efficiently makes ambitious forward-recovery based on “data slices” across versions or data structures both easy to express and efficient. We will show examples that illustrate this flexibility and benefit for sophisticated forward error recovery based on application semantics and application-defined error checking.

GVR requires small source code changes and is efficient. A study with several large applications (OpenMC, preconditioned conjugate gradient (PCG) solver, ddcMD, and Chombo), shows that source code changes are small (<2% LOC), localized, and machine-independent, requiring no software architecture changes. GVR employs a range of techniques to achieve efficient, compact versioning, and measurements show overheads of ~1%. GVR seamlessly exploits SSD and parallel file systems to extend version capacity as appropriate. Overall, our results suggest that GVR's interfaces and implementation are portable, flexible (support a variety of recovery models) and create a gentle-slope path to tolerate growing error rates in future (and current) extreme-scale systems.

Session 5: Future Application Development Environment

Exploiting the User's Knowledge of Resilience

Robert Lucas, *Information Sciences Institute*

By the end of this decade, the Department of Energy (DOE) will be deploying massively parallel systems to address a broad set of problems ranging from mission critical challenges such as nuclear weapons stewardship to addressing fundamental questions in science and technology. These high performance computing systems will be constructed from exascale technology. As such, they will be composed of devices less reliable as those used today, and faults will become the norm, not the exception. This will pose significant problems for DOE users, who for half a century have enjoyed an execution model that relied on correct behavior by the underlying computing system. The University of Southern California (USC), the Lawrence Livermore National Laboratory (LLNL), and the Jet Propulsion Laboratory (JPL) believe that a new generation of dependable applications must be developed to successfully exploit this next generation of technology. Such applications and the systems they run on must be introspective and adaptive, actively searching for errors in their program state with hardware mechanisms and new software techniques. Towards that end, we are extending today's execution model to enable scientific software developers to express their knowledge of resilience in their applications. We are also creating a new Introspection Framework for Resilience that can exploit this knowledge at run time. We are extending the ROSE compiler and the SHINE introspection engine so that faults injected into a resilient application's state will be detected and dealt with, whether by ignoring them, correcting them if possible, or reverting to an earlier checkpoint when necessary.